DESIGN OF NON-LINEAR MULTI-RESOLUTION DECOMPOSITIONS WITH APPLICATIONS TO IMAGE COMPRESSION

Eduardo S. Cardoso Jr., Eduardo A. B. da Silva

PEE/COPPE/DEL/EE, Universidade Federal do Rio de Janeiro
Caixa Postal 68504, Rio de Janeiro, RJ, CEP 21945-970, Brazil
phone: 562-8156, fax:290-6626
{xaua, eduardo}@lps.ufrj.br

Abstract

This paper suggests two novel heuristics to design maximally-decimated non-linear decompositions based on the lifting scheme. The first one dictates the desirable characteristics a non-linear operator should have in order to be able to generate good decompositions. Based on this heuristic, we introduce two non-linear operators using mathematical morphology. The second heuristic is related to the design of good non-linear decompositions using the introduced operators in a lifting scheme. Simulation results show that non-linear decompositions thus generated perform well in image coding schemes.

Introduction

Critically decimated linear multi-resolution decompositions that yield perfect reconstruction play an important role in many signal processing applications, and have been extensively studied [1]. Recently there has been some interest in critically decimated *non-linear* multi-resolution decompositions. In [2], a simple non-linear decomposition scheme was introduced with perfect reconstruction guaranteed irrespective of the decomposition functions. In [3], perfect reconstruction is achieved with more sophisticated structures, by cascading several simple non-linear stages. In [4] and [5], more general structures are presented. A general framework for the theoretical analysis and generalization of such decompositions is presented in [6]. Nevertheless, the problem of designing such decompositions has not yet been solved.

Our main goal is to design and implement decompositions using the structure in figure 1, know as the *lifting* scheme [7]. The operations $T_i(\cdot)$ and $S_i(\cdot)$ are non-linear. One should note that using the analysis structure in figure 1, perfect reconstruction can be achieved by just reverting the order of operations and the signs of the sums in the synthesis operation. The design of such decompositions presents two challenges. One is the definition of the non-linear operations to be used. The other is how to find the m_{t_i} and m_{s_i} such that the resulting decompositions are "good" in some sense.

1. CHOICE OF THE NON-LINEAR OPERATIONS

In several applications, one desires a two-band decomposition such as the one in figure 1 to generate a compact representation of the input signal. This in general implies that, if the polyphase decompositions $\triangle$ and $\bigcirc$ are input to it, the output $\tilde{\triangle}$ is an approximation band and the output $\tilde{\bigcirc}$ is a detail band.

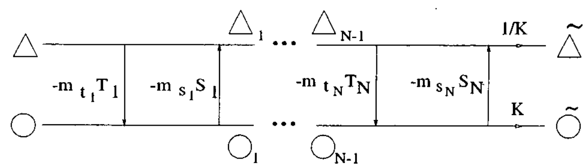

Figure 1: Lifting Structure.

By looking at the first step of lifting in figure 1, we see that the operator $T_1(\cdot)$ generates a prediction of the polyphase component $\bigcirc$ using the polyphase component $\triangle$ to construct the detail band $\bigcirc_1$. Since $\bigcirc_1$ is the result of a subtraction between $\bigcirc$ and its prediction, it is important that the prediction generated by the $T_i(\cdot)$ operator does not add a DC level to its input (one should note that linear operators enjoy this property). Otherwise, the detail band will have a non-zero DC level, which in general leads to non-compact representations. On the other hand, the approximation band is the result of the subtraction between the prediction of the polyphase $\triangle$ (done by $S_1(\cdot)$ using the detail band $\bigcirc_1$) and $\triangle$. Then, for a similar reason, the $S_1(\cdot)$ operator should not add a DC level to its input.

Also, considering that the detail band is "lifted" by the operator $S_i(\cdot)$ to the approximation band, it is convenient that the operators used are smoothing functions. It is interesting to note that several "good" linear decompositions, when decomposed into lifting steps, enjoy this property [7].

The above properties can be summarized as follows:

I. The operators should be smoothing functions, that is, they should generate a lower resolution approximation of the input;

II. They must not add a DC level to its input.

Many non-linear operations enjoy property I, as, for example, *median*, *open* and *close*. However, most of them do not enjoy, in general, property II. For instance, the morphological gray scale operations open, $f \circ k$, and close, $f \bullet k$ do not enjoy property II, because $f \circ k \leq f$ and $f \bullet k \geq f$ [8] [9]. On can show that the same can be said about their compositions $(f \circ k) \bullet k$ and $(f \bullet k) \circ k$. Note that $(f \circ k) \bullet k$ is on average under f and $(f \bullet k) \circ k$ is on average over f.

Fortunately, we can exploit properties of those operations in order to construct a non-linear operation having property II. Remember that $f \circ k \leq f$ and $f \bullet k \geq f$. Therefore, the average of

0-7803-6685-9/01/$10.00©2001 IEEE

those operations will tend to generate an output whose DC level is closer to the one of the input. Thus, a non-linear operation which will enjoy property I and approximately enjoy property II can be as follows:

$$\mathcal{L}_k(f) = \frac{(f \circ k) + (f \bullet k)}{2} \tag{1}$$

Another operator, enjoying prop. II with more accuracy, can be:

$$\mathcal{G}_k(f) = \frac{[(f \circ k) \bullet k] + [(f \bullet k) \circ k]}{2} \tag{2}$$

Simulation results have confirmed that operators $\mathcal{L}_k(\cdot)$ and $\mathcal{G}_k(\cdot)$ satisfy property II reasonably well.

Once a non-linear operator satisfies properties I and II, it still needs to possess one further characteristic, related to the structure in figure 1, in order to generate "good" decompositions. It is explained in what follows.

The equations for stage i of *lifting* are:

$$\triangle_i = \triangle_{i-1} - m_{s_i} S_i(\bigcirc_i) \tag{3}$$

$$\bigcirc_i = \bigcirc_{i-1} - m_{t_i} T_i(\triangle_{i-1}) \tag{4}$$

As mentioned in the previous section, the operator $T_i(\cdot)$ essentially estimates $\bigcirc_{i-1}$ using $\triangle_{i-1}$. On the other hand, $S_i(\cdot)$ modifies $\bigcirc_i$ in order to generate a $\triangle_i$ with "better" characteristics. Remembering that $\bigcirc$ and $\triangle$ are polyphase components of the original signal, we have that $\bigcirc_{i-1}$ is delayed by 1 sample with respect to $\triangle_{i-1}$ in the original signal, and thus by 1/2 sample if one considers the polyphase components as two sub-sampled signals. Then, to perform a good estimate of $\bigcirc_{i-1}$ using $\triangle_{i-1}$, $\triangle_{i-1}$ must be delayed by 1/2 sample. Likewise, since $\bigcirc_i$ is advanced by 1/2 sample with respect to $\triangle_{i-1}$, $S_i(\cdot)$ must advance $\bigcirc_i$ by 1/2 sample in order to modify $\triangle_{i-1}$.

In order to state the above more formally, we define generalized decimation ($\Lambda_\epsilon[.]$) and interpolation ($\Gamma_\epsilon[.]$) operators. These definitions are:

$$\Lambda_\epsilon[f](n) = f(2n+\epsilon) \tag{5}$$

$$\Gamma_\epsilon[f](n) = \begin{cases} f(\frac{n+\epsilon}{2}) & , n+\epsilon \text{ even} \\ \eta & , n+\epsilon \text{ odd} \end{cases} \tag{6}$$

where $\epsilon \in \{0, 1\}$, $n \in \mathbb{N}$ and η is a neutral element.

The neutral element η depends on the operator which will be applied to $\Gamma_\epsilon[f]$. In the linear case, $\eta = 0$ (a FIR filter is a weighted sum). For the morphological gray-scale open, $\eta = +\infty$, and for the morphological gray-scale close, $\eta = -\infty$. This can be easily inferred from the geometrical interpretations of those operations by using the following argument: the gray-scale opening of the signal f with the structuring element k can be seen as the collection of the largest values of the umbra of k such that it still is under f [8]; so, using $\eta = +\infty$ in the interpolation, the open operation will depend only on the values of f. Since the morphological closing is the dual of the opening, analogous considerations imply that, for the closing, $\eta = -\infty$.

Thus, using the above notation, the properties described in above paragraphs are equivalent to saying that $T_i(\cdot)$ and $S_i(\cdot)$ are of the form:

$$T_i(f) = \Lambda_1[\mathcal{F}(\Gamma_0[f])] \tag{7}$$

$$S_i(f) = \Lambda_0[\mathcal{F}(\Gamma_1[f])] \tag{8}$$

where $\mathcal{F}(\cdot)$ is the operator used in the *lifting* steps; for example, it can be either $\mathcal{L}_k(\cdot)$ or $\mathcal{G}_k(\cdot)$ defined in equations 1 and 2. It is important to note that these equations only hold for $\mathcal{F}(\cdot)$ with zero delay; for non-zero delay $\mathcal{F}(\cdot)$, the delay must be compensated for in equations 7 and 8.

It is interesting to note that many popular factorizations of linear FIR filter banks in lifting steps with the method described in [7] implicitly consider relations similar to ones in eqs. 7 and 8. Also, we have verified experimentally that in the non-linear case, this procedure is essential in order to avoid some annoying artifacts introduced by quantization.

2. DESIGN OF CRITICALLY DECIMATED NON-LINEAR DECOMPOSITIONS

Once the $T_i(\cdot)$ and $S_i(\cdot)$ are defined, one naturally wonders what constraints must the scalars m_{t_i} and m_{s_i} in equations 3 and 4 satisfy in order to generate "good" decompositions. For example, in signal compression applications, "good" decompositions are ones that provide good energy compaction properties.

The above concepts can be exemplified through a one stage structure. It outputs the components $\tilde{\triangle}$ and $\tilde{\bigcirc}$ as follows:

$$\tilde{\triangle}_1 = \triangle_0 - m_{s_1} S_1(\bigcirc_0 - m_{t_1} T_1(\triangle_0))$$
$$\tilde{\bigcirc}_1 = \bigcirc_0 - m_{t_1} T_1(\triangle_0) \tag{9}$$

The approach chosen for designing good decompositions was to specify its behavior for the two sequences $\vartheta = \{1, 1, \ldots, 1, 1\}$ and $\varpi = \{1, -1, \ldots, 1, -1\}$ (in linear decompositions they would be equivalent to signals with frequencies 0 and π, respectively).

The operators $T_i(\cdot)$ and $S_i(\cdot)$ implemented as in eqs. 7 and 8 using $\mathcal{L}_k(\cdot)$ or $\mathcal{G}_k(\cdot)$ behave like the identity for constant sequences $\alpha\,\vartheta$, α an scalar, provided that the structuring element k used has symmetry around an extreme value and the axis of symmetry is located at a 1/2 sample position. Referring to figure 2a and recalling the geometrical interpretations of the morphological opening and closing, we see that, with this symmetry, the umbra of k will never fit "inside" the interpolation positions, thus resulting that these positions will be filled with the constant value in its neighboring positions. From figure 2b it is clear that for symmetry around an integer sample position, the operator will not have the above mentioned behavior.

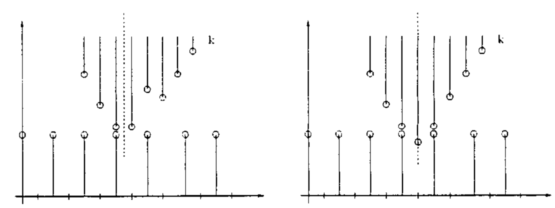

Figure 2: Geometrical interpretation of the closing of a constant sequence: (a) k with desirable symmetry (b) k without desirable symmetry.

Note that both polyphase components of the sequence ϑ are equal to the sequence ϑ itself; the polyphase components of ϖ are the constant sequences ϑ and $-\vartheta$. Then, from eq. 9, the outputs of one lifting stage when those sequences are input are:

- for the ϑ sequence:

$$\tilde{\Delta}_1 = 1 - m_{s_1} \, (1 - m_{t_1}) \quad (10)$$

$$\tilde{\bigcirc}_1 = 1 - m_{t_1} \quad (11)$$

- for the ϖ sequence:

$$\tilde{\Delta}_1 = 1 - m_{s_1} \, (-1 - m_{t_1}) \quad (12)$$

$$\tilde{\bigcirc}_1 = -1 - m_{t_1} \quad (13)$$

Therefore, a good two-channel decomposition would give

- $\tilde{\bigcirc}_1 = 0$, for eq. 11 (in the linear case, it would be analogous to force the DC gain of a high-pass filter to zero);
- $\tilde{\Delta}_1 = 0$, for eq. 12 (in the linear case, it would be analogous to force the gain of a low-pass filter for $\omega = \pi$ to zero).

These two heuristic conditions would uniquely specify m_{t_1} and m_{s_1}.

This procedure can be easily extended for more than one stage; however, since there are only two equations and twice as many unknowns as there are stages, it will generate an infinite number of solutions for the gains m_{t_i} and m_{s_i}.

Experimental results have suggested that the above mentioned restrictions, besides being desirable for a good linear decomposition, also lead to good non-linear decompositions, see table 1. Also, we have generated two-dimensional separable decompositions designed using this heuristic having subbands with directional properties similar to the ones obtained with linear filter banks. This can be seen in figure 3 for the two-stage octave-band decomposition generated by *G0802* defined in section 3.

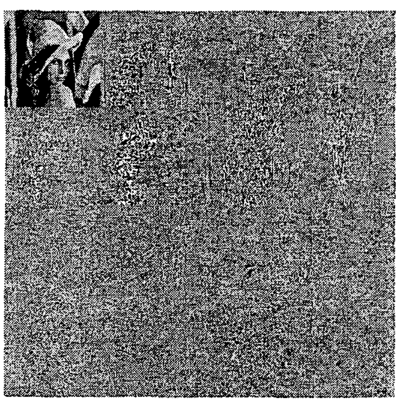

Figure 3: Decomposition for the image Lena 512x512 generated by *G0802* (the detail bands have been multiplied by 4 to enhance visibility)

Since the gain K in figure 1 merely scales the two channels, its value is not relevant as long as restrictions $\tilde{\bigcirc}_1 = 0$ (ϑ input) and $\tilde{\Delta}_1 = 0$ (ϖ input) are concerned. However, one should bear in mind that, since the whole system is non-linear, its value can have an influence on the characteristics of the decomposition. We have verified experimentally that this is indeed the case. For example, if such decompositions are used in image coding schemes, the gain K has a considerable effect on its coding performance.

As a rule of thumb, we have computed K mimicking an important property of wavelet transforms: in eq. 10, $\tilde{\Delta}_1 = \sqrt{2}$.

It is interesting that, for one stage, these heuristics produce the famous Haar wavelets when $\mathcal{F}(\cdot)$ is the identity. The coefficients are $m_{t_1} = 1$ and $m_{s_1} = -0.5$.

3. EXPERIMENTAL RESULTS

In this section, we present three non-linear decompositions designed using the heuristics presented.

The first one is the one stage decomposition with $m_{t_1} = 1$, $m_{s_1} = -0.5$, $\mathcal{F}(\cdot) = \mathcal{G}_k(\cdot)$ and $K = 0.707107$. It was called *Haar-G*.

The second and third ones are 2-stage non-linear decomposition parametrized by $m_{t_1} = 0.8$ and $m_{s_1} = 0.2$, which yields $m_{t_2} = 0.208333$, $m_{s_2} = -0.6528$ and $k = 0.73657$. The second one was implemented using both $\mathcal{F}(\cdot) = \mathcal{G}_k(\cdot)$ and the third one with $\mathcal{F}(\cdot) = \mathcal{L}_k(\cdot)$. They are designated by *G0802* and *L0802*, respectively.

In all decompositions, the structuring element used was a parabola with symmetry around a 1/2 sample position.

For comparison purposes we have employed the non-linear decompositions obtained generating a separable two-dimensional 5-stage decomposition for use in the EZW coder [10]. These results are summarized in table 1, along with results for the Daubechies 9-7 filter bank [11]. We have coded the Lena 512x512 image with a bit-rate of 0.15 bpp.

Filter	Haar-G	G0802	L0802	9-7
PSNR(dB)	29.66	29.92	29.83	31.01

Table 1: PSNR values for the decompositions studied

The reconstructed images for all non-linear decompositions presented have shown a subjective quality comparable to the one of the linear 9-7 wavelet. This can be seen in figures 4 (a, b, c and d), where a zoom of its face was made to highlight the artifacts. With a careful examination, we can see that the non-linear wavelets generate a little less ringing artifacts, together with a slight increase in blockiness.

4. CONCLUSION

In this paper we introduced two new non-linear operations based on mathematical morphology. Their development was motivated by the identification of properties an operator should have in order to be used in critically decimated non-linear multi-resolution decompositions based on the lifting scheme.

We have also introduced a heuristic approach to the problem of designing such decompositions. Three decompositions were designed and tested using the EZW coder, generating reconstructed images with a subjective quality comparable to the best linear ones in the literature. However, an important point is that this heuristics leads to several different solutions with different coding performances. Therefore, we still need criteria for finding the best among such solutions. Nevertheless, considering the flexibility provided by non-linear decompositions, this heuristic stands as a powerful tool for designing alternative signal decompositions.

Further flexibility can be obtained by extending this work to multi-dimensional non-separable decompositions, a subject that will be dealt with in another publication.

It is important to note that the design of non-linear maximally decimated decompositions is still an open problem. However, we

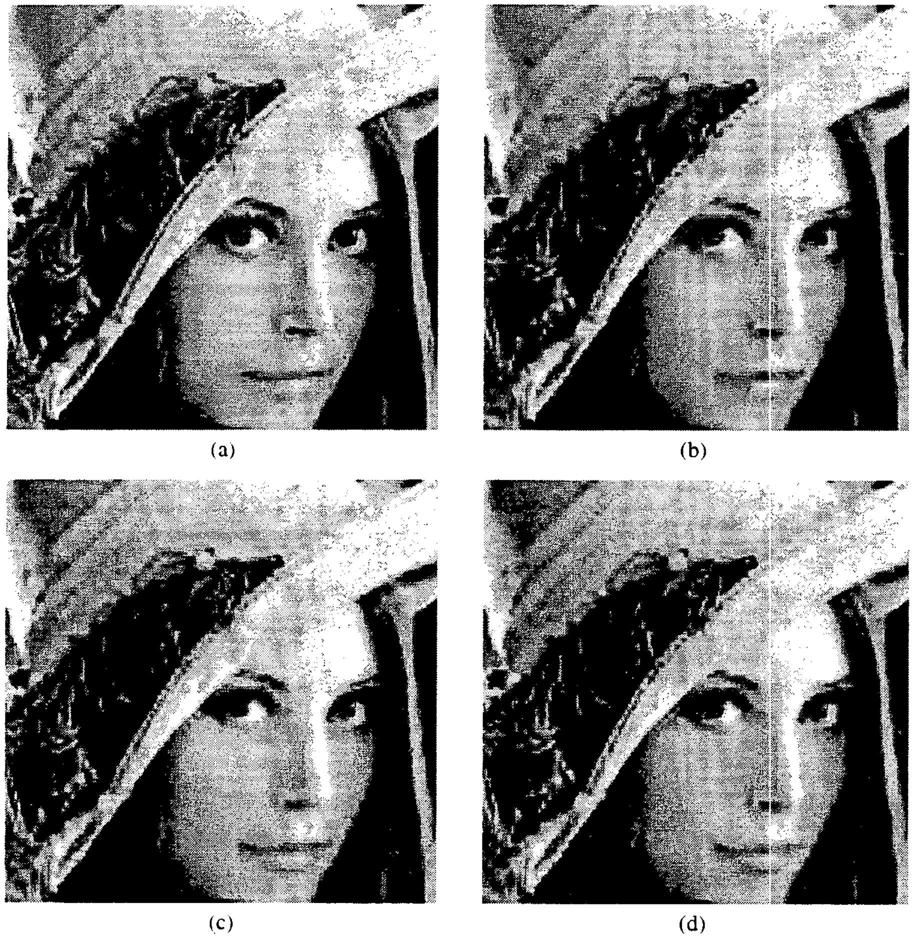

Figure 4: Zoom of Lena 512x512 at 0.15bpp for: (a) 9-7 (b) *Haar-G* (c) *G0802* (d) *L0802*

believe that, although we have presented just a heuristic approach to such design, its strength lies in the fact that it opens new research possibilities in the direction of finding more effective design methods.

5. REFERENCES

[1] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[2] W. L. O. Egger and M. Kunt, "High compression image coding using an adaptive morphological subband decomposition," in *Proceedings of IEEE*, vol. 83, pp. 272–287, 1995.

[3] D. A. F. Florêncio and R. W. Schafer, "Perfect resconstructing nonlinear filter banks," in *Proceedings of the 1996 ICASSP Conference*, vol. III, pp. 1815–1818, 1996.

[4] F. J. Hampson and J.-C. Pesquet, "A nonlinear subband decomposition with perfect reconstruction.," in *Proceedings of the 1996 ICASSP Conference*, vol. III, pp. 1523–1526, 1996.

[5] F. J. Hampson and J.-C. Pesquet, "M-band nonlinear subband decompositions with perfect reconstruction," *IEEE Transactions on Image Processing*, vol. 7, pp. 1547–1560, november 98.

[6] J. Goutsias and H. J. A. M. Heijmans, "Multiresolution signal decompositions schemes. part 2: Morphological wavelets.," Technical Report PNA-R9905, CWI, Amsterdam, The Netherlands, June 1999.

[7] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps.," tech. rep., Bell Laboratories - Lucent Technologies, 1996.

[8] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*. Addison-Wesley Publishing Company, 1992.

[9] J. Serra, *Image Analysis and Mathematical Morphology*. Academic Press, 1982.

[10] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 41, pp. 3445–3462, December 1993.

[11] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Transactions on Image Processing*, vol. 1, pp. 205–220, April 1992.